Advanced Econometrics

Lecture 11: Resampling Methods

Instructor: Ma, Jun

Renmin University of China

December 25, 2021

Asymptotic normality

- In previous lectures, we have so many estimators with the property
$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) \to_d \mathrm{N}\left(0, \sigma^2\right)$$
and equivalently we can write $\widehat{\theta}_n \overset{a}{\sim} \mathrm{N}\left(\theta, \sigma^2/n\right)$.
- Once we have a consistent estimator $\widehat{\sigma}_n$ of σ, the standard error is defined to be $SE = \widehat{\sigma}/\sqrt{n}$. A confidence interval with approximate 95% coverage probability is $\left[\widehat{\theta}_n \pm 1.96 \times SE\right]$.
- We use $\mathrm{N}\left(\theta, \sigma^2/n\right)$ as approximation to the unknown true (often called finite-sample) distribution of $\widehat{\theta}_n$.
- To estimate σ^2 based on the analogue principle (i.e., replace population moments/unknown quantities by their sample moments/estimates), we need knowledge of the expression(formula) of σ^2. Very often the expression is very complicated.
- There are two computation-intensive approaches that do the estimation without requiring knowledge of the expression of σ^2.

Jackknife standard errors

- Suppose our data is $(Y_i, X_{i1}, ..., X_{ik})$ for $i = 1, ..., n$. Denote $Z_i = (Y_i, X_{i1}, ..., X_{ik})$.
- Suppose the estimator $\widehat{\theta}$ can be written as $\widehat{\theta}_n = \varphi_n(Z_1, ..., Z_n)$, e.g., $\varphi_n(z_1, ..., z_n) = n^{-1}\sum_{i=1}^{n} z_i$.
- Now denote $\widehat{\theta}_{-j} = \varphi_{n-1}(Z_1, ...Z_{j-1}, Z_{j+1}, ..., Z_n)$, i.e., $\widehat{\theta}_{-j}$ is an estimator obtained by removing the j-th observation from the entire sample. The variation in $\left\{\widehat{\theta}_{-j} : j = 1, ..., n\right\}$ should be informative about the population variance of $\widehat{\theta}_n$.
- Denote $\overline{\widehat{\theta}} = n^{-1}\sum_{j=1}^{n}\widehat{\theta}_{-j}$. The Jackknife standard error is

$$\widehat{se}_{jk} = \sqrt{\frac{n-1}{n}\sum_{j=1}^{n}\left(\widehat{\theta}_{-j} - \overline{\widehat{\theta}}\right)^2}.$$

- A 95% confidence interval is

$$\left[\widehat{\theta}_n - 1.96 \cdot \widehat{se}_{jk}, \widehat{\theta}_n + 1.96 \cdot \widehat{se}_{jk}\right].$$

- Indeed one can show

$$(n-1)\sum_{j=1}^{n}\left(\widehat{\theta}_{-j}-\overline{\widehat{\theta}}\right)^2 \to_p \sigma^2.$$

- Consider the following simple example: for i.i.d. random variables $X_1, ..., X_n$, we use the sample average $\overline{X}$ as an estimator of $\mu = \mathrm{E}\left[X_1\right]$. It is known that $\sqrt{n}\left(\overline{X}-\mu\right) \to_d \mathrm{N}\left(0,\sigma^2\right)$, where $\sigma^2 = \mathrm{Var}\left(X_1\right)$ in this case.
- For this case,

$$\widehat{\theta}_{-j} = \frac{1}{n-1}\left(n\overline{X}-X_j\right),$$

$$\begin{aligned}\frac{1}{n}\sum_{j=1}^{n}\widehat{\theta}_{-j} =& \frac{1}{n\left(n-1\right)}\sum_{j=1}^{n}\left(n\overline{X}-X_j\right)\\ =&\overline{X}.\end{aligned}$$

- For this simple case,

$$\widehat{\theta}_{-j} - \overline{\widehat{\theta}} = \frac{1}{n-1}\left(n\overline{X} - X_j\right) - \overline{X} = \frac{1}{n-1}\left(\overline{X} - X_j\right).$$

- We have

$$(n-1)\sum_{j=1}^{n}\left(\widehat{\theta}_{-j} - \overline{\widehat{\theta}}\right)^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_j - \overline{X}\right)^2,$$

which is the sample variance that is a consistent and unbiased estimator for σ^2.

Bootstrap

- The bootstrap takes the sample (the values of the realized explanatory and explained variables) as the population.
- The bootstrap is an alternative way to produce approximations for the true distribution of $\widehat{\theta}_n$.
- Note that both asymptotic theory and the bootstrap only provide approximations for finite-sample properties.
- A bootstrap sample is obtained by independently drawing n observations from the observed sample $\{Z_i\}_{i=1}^n$ with replacement.
- The bootstrap sample has the same number of observations as the original sample, however some observations appear several times and others never.

Bootstrap Standard Errors

- Step 1: Draw B independent bootstrap samples. B can be as large as possible. We can take $B = 1000$.
- Step 2: Estimate θ with each of the bootstrap samples, $\widehat{\theta}_b^*$ for $b = 1, ..., B$.
- Step 3: Estimate the standard deviation of $\widehat{\theta}$ by

$$\widehat{se}_{bs} = \sqrt{\frac{1}{B} \sum_{b=1}^{B} \left(\widehat{\theta}_b^* - \widehat{\theta}^* \right)^2}$$

  where $\widehat{\theta}^* = B^{-1} \sum_{b=1}^{B} \widehat{\theta}_b^*$.
- Step 4: The bootstrap standard errors can be used to construct approximate confidence intervals and to perform asymptotic tests based on the normal distribution, e.g. if the coverage probability is 95%, a 95% confidence interval is $\left[\widehat{\theta}_n - 1.96 \cdot \widehat{se}_{bs}, \widehat{\theta}_n + 1.96 \cdot \widehat{se}_{bs}\right]$.

Bootstrap percentile confidence intervals

- Step 1: Draw B independent bootstrap samples. B can be as large as possible. We can take $B = 1000$.
- Step 2: Estimate θ with each of the bootstrap samples, $\widehat{\theta}_b^*$ for $b = 1, ..., B$.
- Step 3: Order the bootstrap replications such that

$$\widehat{\theta}^*_{(1)} \leq \widehat{\theta}^*_{(2)} \leq \cdots \leq \widehat{\theta}^*_{(B)}.$$

- Step 4: The lower and upper confidence bounds are $B \times (\alpha/2)$-th and $B \times (1 - \alpha/2)$-th ordered elements. For $B = 1000$ and $\alpha = 5\%$, these are the 25th and 975th ordered elements. The estimated $1 - \alpha$ confidence interval is $\left[\widehat{\theta}^*_{(B\times(\alpha/2))}, \widehat{\theta}^*_{(B\times(1-\alpha/2))}\right]$.
- Bootstrap percentile confidence intervals often have more accurate coverage probabilities (i.e. closer to the nominal coverage probability $1 - \alpha$) than the usual confidence intervals based on standard normal quantiles and estimated variance.

Bootstrap-t test

- We consider testing $\mathrm{H}_0 : \theta = \theta_0$.
- We can conduct a bootstrap-based hypothesis testing based on the bootstrap percentile confidence interval: we simply reject H_0 if θ_0 fails to be an element of the bootstrap percentile confidence interval.
- We can show that $T = \sqrt{n}\left(\widehat{\theta} - \theta_0\right)/\widehat{\sigma} \rightarrow_d \mathrm{N}(0,1)$ under H_0. We use the standard normal distribution as approximation to the true distribution of T and define critical values based on standard normal quantile.
- For each bootstrap sample $b = 1, ..., B$, we can calculate $\widehat{\sigma}^*$ using the bootstrap sample.

- Step 1: Draw B independent bootstrap samples. B can be as large as possible. We can take $B = 1000.$
- Step 2: Estimate θ and σ with each of the bootstrap samples, $\widehat{\theta}_b^*$, $\widehat{\sigma}_b^*$ for $b = 1, ..., B$ and the t-value for each bootstrap sample:

$$t_b^* = \frac{\sqrt{n}\left(\widehat{\theta}_b^* - \widehat{\theta}\right)}{\widehat{\sigma}_b^*}.$$

- Step 3: Order the bootstrap replications of t such that $t_{(1)}^* \leq t_{(2)}^* \leq \cdots \leq t_{(B)}^*$. The lower critical value and the upper critical value are then the $B \times (\alpha/2)$-th and $B \times (1 - \alpha/2)$-th ordered elements. For $B = 1000$ and $\alpha = 5\%$, these are the 25th and 975th ordered elements. The bootstrap lower and upper critical values generally differ in absolute values.

- A common mistake is that in Step 2, one mistakenly computes

$$\frac{\sqrt{n}\left(\widehat{\theta}_n^{*(b)} - \theta_0\right)}{\widehat{\sigma}_n^{*(b)}}.$$

  The test will have no power if we made this mistake.

- The distribution of the t-statistic $T = \frac{\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right)}{\widehat{\sigma}_n}$ under $\mathrm{H}_1 : \theta \neq \theta_0$ is different from that under H_0. Under H_1, T is not centered:

$$T = \frac{\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right)}{\widehat{\sigma}_n} = \frac{\sqrt{n}\left(\widehat{\theta}_n - \theta\right)}{\widehat{\sigma}_n} + \frac{\sqrt{n}\left(\theta - \theta_0\right)}{\widehat{\sigma}_n}.$$

- An important guideline is that we should always approximate the distribution of T under H_0, i.e., the distribution of $\sqrt{n}\left(\widehat{\theta}_n - \theta\right)/\widehat{\sigma}_n$.

- In finite samples (fixed n), for neither the bootstrap-t test nor the usual t-test that uses ± 1.96 as critical values , the true probability of making type-I error is exactly equal to α (e.g., 0.05).
- In almost all cases, the true probability of making type-I error is greater than α, i.e., we always "over-reject" the null hypothesis.
- One can show that for bootstrap-t test, in finite samples, the true probability of making type-I error is closer to the nominal significance level α than the standard t-test that uses ± 1.96 as critical values.